ARL-TN-0708 • Oct 2015

US Army Research Laboratory

Potential Technologies for Assessing Risk Associated with a Mesoscale Forecast

by Patrick A Haines, Jeffrey A Smith, Mark R Hjelmfelt, William J Capehart, and James L Cogan

NOTICES

Disclaimers

ARL-TN-0708 ● OCT 2015

US Army Research Laboratory

Potential Technologies for Assessing Risk Associated with a Mesoscale Forecast

by Patrick A Haines, Jeffrey A Smith, and James L Cogan
Computational and Information Sciences Directorate, ARL

Mark R Hjelmfelt and William J Capehart
Atmospheric Sciences Department, South Dakota School of Mines and Technology

REPORT DOCUMENTATION PAGE						*Form Approved* *OMB No. 0704-0188*
Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing the burden, to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. **PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.**						
1. REPORT DATE *(DD-MM-YYYY)* October 2015	**2. REPORT TYPE** Final					**3. DATES COVERED (From - To)** 01 Oct 2014–30 Sep 2015
4. TITLE AND SUBTITLE Potential Technologies for Assessing Risk Associated with a Mesoscale Forecast						**5a. CONTRACT NUMBER**
						5b. GRANT NUMBER
						5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S) Patrick A Haines, Jeffrey A Smith, Mark R Hjelmfelt, William J Capehart, and James L Cogan						**5d. PROJECT NUMBER**
						5e. TASK NUMBER
						5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) US Army Research Laboratory ATTN: RDRL-CIE-M White Sands Missile Range, NM 88002						**8. PERFORMING ORGANIZATION REPORT NUMBER** ARL-TN-0708
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)						**10. SPONSOR/MONITOR'S ACRONYM(S)**
						11. SPONSOR/MONITOR'S REPORT NUMBER(S)
12. DISTRIBUTION/AVAILABILITY STATEMENT Approved for public release; distribution is unlimited.						
13. SUPPLEMENTARY NOTES						
14. ABSTRACT Warfighters face uncertainty in every task undertaken during mission execution. Military decision-making seeks to reduce these uncertainties through planning contingencies whenever possible. Information can reduce uncertainty—but only if the information is understood in context; that is, if warfighters can evaluate the information in light of their plans and contingencies to assess mission-related risks. Information context is extremely important to the warfighter, and that context is a 4-dimensional (4D)—x, y, x, t—cube. Numerical Weather Prediction (NWP) forecasts produce 4D information, yet the means to interpret it in warfighter-relevant terms is not easily available. It is not currently possible to anticipate the meteorological errors of a particular forecast. The risk in using a meteorological forecast of any kind is that it will inaccurately predict the future state of the atmosphere to an unknown extent. Inaccuracy is due to errors in the initial conditions and errors introduced by the numerical representation and physical parameterizations of an NWP forecast model. Despite these limitations, some forecasts are quite accurate and use of them in the planning and execution of operations would be highly beneficial. The salient question: Can the quality of the currently available forecast be known without waiting for confirming observations?						
15. SUBJECT TERMS Numerical forecast, risk, uncertainty						
16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**		**19a. NAME OF RESPONSIBLE PERSON** Patrick A Haines
a. REPORT Unclassified	**b. ABSTRACT** Unclassified	**c. THIS PAGE** Unclassified	UU	24		**19b. TELEPHONE NUMBER (Include area code)** 575-678-5592

Standard Form 298 (Rev. 8/98)
Prescribed by ANSI Std. Z39.18

Contents

List of Tables

1. Introduction

Warfighters face uncertainty in the performance of each and every task they undertake during mission execution. The military decision-making process seeks to reduce these uncertainties through planning by creating contingencies whenever possible. Information can reduce uncertainty; however, this reduction can only happen if the information is understood in context, that is, if the warfighters can evaluate the gained information in light of their plans and contingencies to assess mission-related risks. Information context is extremely important to the warfighter, and that context is a 4-dimensional (4D)—x, y, x, t—cube. Numerical Weather Prediction (NWP) forecasts produce 4D meteorological information yet the means to interpret the data cube in warfighter-relevant terms is not easily available. NWP forecasts have improved greatly since NWP's inception in 1950 but remain imperfect. It is not currently possible to anticipate ahead of time the meteorological errors of a particular forecast.

The risk in using a meteorological forecast of any kind is that it will inaccurately predict the future state of the atmosphere to an unknown extent. There are 3 reasons for forecast inaccuracy:

1) The forecast model's initial conditions contain error due to an observation system that cannot adequately sample the complexity of atmospheric conditions at the start time of a forecast. In addition, the observations on which the initial conditions are based contain some amount of error that may be unavoidable.

2) Although the forecast model is based on continuous partial differential equations, the actual forecast process is carried out with discrete or truncated numerical approximations of these partial differential equations. These approximations induce error due to their discretization and present a minimum scale at which physical features (e.g., complex terrain) and atmospheric processes can be adequately represented in an NWP forecast.

3) As a consequence of 2) above, the forecast models have to parameterize atmospheric physical processes (e.g., radiative transfer, precipitation and evaporation, surface heating, turbulence, and so on) that cannot be explicitly represented in order to account for their effect on the evolution of the atmosphere. These parameterizations can be extremely complex; nonetheless, they imperfectly represent the actual physics.

Despite the above limitations of NWP, some forecasts are quite accurate and the use of them in the planning and execution of operations would be highly beneficial.

However, other forecasts are less accurate and the use of them in planning and execution could adversely impact the operations to the point of causing failure and/or unnecessary loss of equipment and lives. The salient question is this: Can the quality of the currently available forecast be known without waiting for confirming observations?

In this report we seek to assess 2 potential methods to ascertain the quality of an NWP forecast: ensemble modeling and use of a qualitative "confidence index" (CI). The development and test of one or both of these methods for possible operational use will require significant resources and is not a short-term effort—but, if successful it could lead to a major improvement in weather support for planning and mission execution.

2. Overview of Assessment Methods

There are currently 2 ways to assess the quality of a current forecast:

1) Ensemble modeling, and
2) Qualitatively, based on forecasting experience and the accuracy of the most recent forecasts; this empirically based method has been quantified into calculation of a CI.

Ensemble modeling involves starting the same forecast model with slightly different initial conditions, or starting a variety of models with essentially the same initial conditions and evaluating the evolution of the different forecasts. This can be a measure of the forecast uncertainty if the breadth of the initial conditions or variety of models used captures the true atmospheric evolution during the forecast period. The ensemble forecast(s) can also be used to produce an improved forecast.

Ensemble modeling also can be accomplished through the use of different physical parameterizations in a forecast model with the same or varied initial conditions. As with the different initial conditions, the set of different physical parameterizations as a whole must capture the actual physics. Ideally, the ensembles would include both variation in initial conditions and physical parameterizations; however, the number of ensemble members might then be unfeasibly large. Ensemble modeling experience has shown that variation in initial conditions is more important than variation of physical parameterizations to forecast quality.[1]

With initial conditions being the more important reason for forecast uncertainty, there appears to be a viable alternative to ensemble modeling. This is based on many forecasters' experience with a large set of initial conditions, forecasts, and the subsequent weather that occurred. The CI is a rules-based assessment

concerning the initial condition that are associated with better or poorer forecast outcomes[2] show that a "trained" rule set can explain a majority of the variance in subsequent atmospheric evolution based on application of the trained rules set to initial conditions. The CI requires a great deal of preliminary training but, once trained, can be quickly applied to any appropriate set of current conditions resulting in a numerical value indicating the confidence that a forecast based on those conditions will be good or not. It may also be able to provide quantitative expectations about the forecast meteorological (MET) parameter accuracy such as the mean absolute errors of temperature, pressure, and wind that are important to artillery accuracy.

3. Assessment of Spatial and Temporal Uncertainties in Ensemble Forecasts

One can view the forecast produced by an NWP as a single-sample estimate of the future atmospheric state described over a 4D region of space. Ensemble prediction methods, in one form or another, generate a "large" sample of the future state of the atmosphere; thus, they provide a means to estimate both the atmospheric state, as described by conditions at a given point in space, as well as the uncertainties, expressed as the range of possible states. For example, possible conditions could include extreme winds or heavy rains. By themselves, some possible states have informational value (i.e., one cannot fly certain unmanned aerial vehicles in high-wind conditions) but the more-desired information for a warfighter would answer the question, "How likely is that condition to occur over a given region of the battlespace?"

To help address this question, we can employ a Geographic Information System (GIS) in conjunction with high-resolution, Army-scale NWP codes executed over tactical-sized domains. A GIS provides a single framework within which we can easily incorporate available data products such as digital elevation models and land-use characteristics as well as the ensemble forecasts for analysis.[3] Within the GIS, we can estimate uncertainties from the ensemble forecasts produced via NWP, for example, over complex terrain where weather phenomena are particularly influenced by terrain geometry and the variance of heat momentum and moisture fluxes resulting from the terrain characteristics.

The objectives of the analysis will be to 1) quantify and assess the accuracy the spatial and temporal uncertainties present in ensemble forecasts, and 2) develop methods to depict MET uncertainty in ways that enable warfighters to quickly assess operational risks and exploit favorable weather conditions.

3.1 Technical Approach

There are 3 forms of uncertainty that come with the use of NWP to produce forecasts: 1) the input uncertainty that derives from an incomplete observation of the current atmospheric state, 2) the model uncertainty that is a consequence of an incomplete understanding or description of the atmospheric physics, and 3) the forecast uncertainty; that is, the likelihood that a single prediction is correct. In order to properly understand the forecast uncertainty, we must understand how input and model uncertainties influence it.

1) Develop uncertainty measures:
   a. Characterize input uncertainty by subsampling initialization data to determine sensitivity to observation density, region, and synoptic conditions.
   b. Characterize model uncertainty by evaluating the range of possible parameterizations.
   c. Characterize forecast uncertainty by binning the output space and estimating bin probability.

2) Define the experiment: Employ statistical design of experiments using a range of regions, synoptic conditions, parameterization schemes, and input densities to generate a range of ensemble forecasts generated under very controlled conditions. This step will allow us to understand how input and model uncertainties drive output or forecast uncertainties as well as determining a NWP configuration with broad applicability.

3) Evaluate the experimental results:
   a. Evaluate the ensemble forecast by comparing the spread with that of weather observations, by determining the accuracy of the probabilistic forecasts, and by verifying the accuracy of the mean prediction using traditional and location-based GIS approaches.
   b. Extract the probability distribution function (PDF) of the frequency distribution of the variables of interest.
   c. Use the PDF to interpret the forecast in terms of the probability of an event or the expected spread or variance of variables produced by the NWP forecast occurring at a particular point and time.
   d. Combine the probability of the event with the deterministic forecast of the event into a merged product, which expresses the confidence of the event.

4) Develop visualization methods:
   a. Depict the merged products for a specified event in terms of the certainty of a specified forecast value (e.g., region of highest snowfall

superimposed over region of highest certainty to reveal the area of highest expected snowfall).

b. Develop methods to depict merged products to enable warfighters to quickly assess the risks involved in conducting the mission using the GIS capability for visualization of complex information. A candidate method using ArcGIS[4] software uses a 3-color technique to visualize the information in a continuous fashion, pixel-by-pixel, that can be more meaningful and easier to interpret. Through use of combinations of red, green, and blue lights the information can be displayed which is easily visible to the human eye and makes the information more understandable and more easily determined from the forecast.

c. Tailor the final visualization product to maximize the effectiveness for use by warfighters by applying developed sets of criteria including:
   1) Resolution in time and space matches the scale of the operational activities.
   2) Lead time matches the decision horizon.
   3) Meteorological variables cover the warfighter scope of interest.
   4) Forecast uncertainty is minimized.
   5) Predicted probabilities of discreet events correspond to verified frequencies of occurrence.

4. Assessing the Value of CI

The CI is a metric of inherent risk of a meteorological situation not being adequately predicted by a forecast model. It uses a combination of intrinsic (non-model related) meteorological fields and recent model performance and in practice consists of a rule set evaluation of synoptic scale features. The NATO Military Meteorology panel has extensively reviewed CI and found that it is a valid way to assess forecast uncertainty, but also thought at that time that quantifying uncertainty on the basis of ensemble model results would in the long run offer more on this problem than CI.

In a more recent review, however, CI showed more promise. It had been further developed so that its rule set was evaluated on the basis of one year's Global Forecast System (GFS) data. Some rules were discarded and the CI score was now based on a summation of the rules in which the individual rules are each weighted by a coefficient. The coefficients were derived by multiple linear regression for a set of 2,000 × 2,000-km domains using initially the one year of GFS data. The domains cover all global land areas and the coefficients differ by domain. Even more recently several years of data have been used.

CI is based on the idea that there are identifiable meteorological features that are indicative of variable, unstable or otherwise low predictive states that may reduce human and numerical forecast skill. Some examples of challenging features are new fronts or fronts that change speed (or start to move); formation of low pressure areas (cyclones); upper-level disturbances; and organized convection. These complicating features impart a decrease in “confidence” in the verbatim interpretation of the model-predicted forecast.

The current version of CI (CI v0.3b) is applicable to 12- and 24-hour (h) Meteorological Gridded Message (METGM) forecasts for any 2,000 × 2,000-km land domain for the full latitude range of valid METGMs. The CI provides an estimate of the uncertainty of the forecast which has been related to the standard deviation of the error for meteorological variables at various height levels. CI has been formally tested with the Canadian Global Environmental Multiscale model and the American GFS models, and informally applied on the Weather Research and Forecasting (WRF) model.

The current CI equation is as follows:

$$
\begin{aligned}
&\mathrm{FIELD\_DIFF\_VER} = b0 + b1\ \mathrm{GRAD\_MAX} + b2\ \mathrm{GRAD\_MIN} + b3 \\
&\mathrm{MAX\_CLOSEDLO} + b4\ \mathrm{FIELD\_DIFF\_CON} + b5\ \mathrm{MAX\_RV} + b6 \\
&\mathrm{MAX\_RV\_SIZE} + b7\ \mathrm{MAX\_MAGGEO} + b8\ \mathrm{MAX\_MAGGEO\_SIZE} + b9 \\
&\mathrm{SIG\_MAGGEO\_SIZE} + b10\ \mathrm{MAX\_MAG\underline{A}GEO} + b11 \\
&\mathrm{MAX\_MAG\underline{A}GEO\_SIZE} + b12\mathrm{SIG\_MAG\underline{A}GEO\_SIZE} + \mathrm{b13CAPE\_MAX} \\
&+ \mathrm{b14CIN\_MAX} + \mathrm{b15DIV\_SIZE\_MAX} + \mathrm{b16DIV\_Maximum} + \\
&\mathrm{b17CONV\_SIZE\_MAX} + \mathrm{b18CONV\_Maximum}
\end{aligned}
$$

The 18 predictors in this equation are defined in Appendix B.[2] This CI equation is a linear equation with 19 undetermined coefficients, each corresponding to a specific meteorological feature of importance. The values corresponding to each of the 19 coefficients were determined for each domain by performing a multiple linear regression for all of either 12-h or 24-h forecasts for almost an entire year of data (02 May 2008 to 30 Apr 2009 and initiated at 00 and 12 coordinated universal time) giving about 600 valid data points. The regressions are calculated using the “R” software package.

The CI v0.3b was trained for a global set of domains on a 1-year data set of 12- and 24-h forecasts. Individually and collectively, both 12- and 24-h forecasts for the 119 domains fell within the same envelope of CI_RAW score versus the 500 hectopascals (hPa) forecast height. For even the worst-performing Antarctic domains, less than 5% of the points fell outside this envelope. Training on 12-h forecasts produced similar distribution of forecast-error points as a function of

CI_RAW score as the 24-h forecasts within the same envelope, but the points were concentrated at lower model-forecast errors and CI_Risk scores. In most individual cases the model forecast errors were smaller for 12-h forecasts than for the same forecasts carried on to 24 h; however, for a small percent of cases the 12-h forecast errors were larger. The largest forecast errors and CI_RAW scores were found at the highest latitudes; the smallest forecast errors and CI_RAW scores were found near the equator. Generally, the highest correlations and best fit to the CI conceptual model occurred in well-developed regions (with lots of observations) in midlatitudes.

To run CI operationally over a domain that does not coincide with one of the 119 trained domains, CI coefficients obtained from the nearest trained neighbor domain are used. To evaluate CI performance in operational mode, CI was run operationally over the nearly one year of data set for 14 selected domains. CI was rerun in training mode over the same domains and the results compared. For some domains the penalty for using the nearest neighbor's coefficients was small. For others it was larger and increased as a function of distance to the neighbor. However, it is concluded that CI performs adequately in an operational mode.

It is important that CI be exercised for different time periods and models to determine the robustness of these results. It is highly recommended that additional training be undertaken with domains placed at a higher density, especially in latitude, than used in this case.

To develop further, CI should be applied to different forecast models and additional years. In addition, these future CI versions should be trained on overlapping domains in order to improve the accuracy when the nearest neighbor coefficients are used in an operational mode.

The separate calculations of Convective Available Potential Energy and Convective Inhibition should be replaced with a Convection Index; this should improve the usefulness of the combined rule. Other rules involving heights at pressure levels different than 500 hPa (e.g., 850 hPa and 300 hPa) should be considered. This is because atmospheric jets typically are found in the upper troposphere and convergence and advection in the lower troposphere is often important. Evaluating these only at 500 hPa misses their greatest impact.

There are potential users of CI beyond the ballistics community; they may be primarily interested in different variables or atmospheric levels, so a different rule set may be necessary to meet their needs. Ensemble forecasts are becoming more widely available; further development should consider the roles of ensembles, CI, and ensembles as a part of CI. Note that the variance of the ensembles is not the same as the variance of the atmosphere, so a calibration such as that performed for

CI v0.3b will have to be done (and ensemble variance properties can be very ensemble dependent).

CI has been proven to be a viable way to assess whether a forecast will be good because a high CI score has shown that such a forecast will be good. On the other hand, a forecast with a low CI score can turn out good or bad. Therefore, the risk of using a high CI forecast is low and the risk of using a low CI forecast is much higher. By implementing new rules (as mentioned previously) and extending the training to additional models and years of data, the spread of model-forecast variation with CI score is expected to be reduced. This will enhance the value of CI. In the results so far, there has been general agreement between the CI score and the actual model-forecast error—the 500-hPa root-mean-square error (RMSE) height error. Future work should also consider other forecast metrics such as the 300- and 850-hPa RMSE–vector wind errors, the 850-hPa RMSE temperature and height errors, and so on.

5. Conclusions

This brief report presents 2 methodologies, ensemble modeling and the Confidence Index, which potentially could produce a means to estimate the risk associated with a mesoscale forecast. A number of forecast centers nationally and internationally such as the National Center for Environmental Prediction, the US Air Force 557th Weather Wing, and the European Center for Medium-Range Weather Forecasts produce global forecasts using ensemble methods. Nevertheless, the computational load may be excessive for smaller operational computing systems, especially for larger model domains. However, as higher-end computing capabilities become available in smaller systems over the next several years, a useful ensemble method may become feasible on a field computer. CI methods have shown promise and could run on smaller systems quickly enough to be useful once the appropriate training has been completed for the selected regions. Further development of the CI methodology is needed before being testing and released for operational applications. The potential exists for a method that would combine ensemble and CI algorithms, though, that would require additional investigation.

Either methodology or a combination of the 2 potentially could provide important added value to meteorological support by providing users with an indication of the quality of a mesoscale forecast and consequently applications that depend on those forecasts. However, development of a software package for operational use will require significant resources and effort. It is not short term, but interim methods along the way could indicate the way forward for development of an operational package.

6. References and Notes

1. Personal communication, Richard Penc with Patrick Haines. 2015 Sep.

2. Hjelmfelt MR, Capehart WJ, Farley RD. Training evaluation technical report. Atmospheric Sciences Department, South Dakota School of Mines and Technology, Rapid City (SD); 2013 Oct. SDSMT-ARDEC Contract W15QKN-06-D-0006 0010 TASK 3.1.1, Training and Evaluation Document CI V0.3b.

3. Foley T, Smith JA, Raby JW, Reen B, Penc R. Investigating surface bias errors in the Weather Research and Forecasting (WRF) model using a Geographic Information System (GIS). J App Met Climat. Forthcoming 2016.

4. ArcGIS for desktop with ArcGIS geostatistical analysis and spatial statistics toolboxes. Environmental Systems Research Institute (ESRI), Inc. 2015.

INTENTIONALLY LEFT BLANK.

Appendix A. Aspects of Numerical Weather Prediction

Numerical Weather Prediction (NWP) is the science of forecasting weather conditions at some time in the future based on present and past observations using complex mathematical/physical model(s) and advanced computational techniques. These predictions can range from hours to months and in some cases years into the future. Spatially, these codes can simulate conditions from small regional microclimates to global climates. Of particular interest to the US Army are codes that can make short-term predictions that are valid for a few hours over a roughly brigade-sized area of responsibility.

While both long- and short-term forecasts are of interest to a brigade commander for planning purposes, it is at the meso- and microscales where forecasts can enhance situational awareness for the echelons below brigade. Mesoscale models resolve meteorological features such as thunderstorms or microbursts that typically are on the order of fewer than 10 kilometers (km) in extent. Microscale models resolve features are often less than 1 km in extent, such as turbulent flows in an urban canyon. Temporally, these meso- and microscale models provide needed detail that is not provided by the coarse synoptic forecasts of large-scale features, such as weather fronts that update typically every 6 to 12 hours (h). For US Army purposes, the term "Nowcast" refers to forecasts of mesoscale conditions up to 6 h into the future and can be updated as frequently as every 30–60 minutes.

Nowcast models are typically high-resolution applications of research and operational mesoscale NWP models such as the Advanced Research Weather Research and Forecast model (WRF-ARW) with extensions that might include finer terrain resolutions and more detailed representations of the underlying atmospheric physics. Functionally, these models take as initial conditions both very coarse-grained synoptic forecasts and observational data from various meteorological data sources to produce their higher resolution Nowcasts.

The foundation of NWP models is the conservation of mass, heat, motion, and water vapor as well as other gaseous and aerosol materials over the region of interest. NWP codes model these properties through a set of coupled partial differential equations with the first-guess–initial and time-dependent–later boundary tendencies provided by a forecast from a coarse-grained synoptic model. Because these models seek to represent climate physics at a high resolution, the transition from coarse- to fine-grain is handled via a nesting strategy so that boundary conditions enter into the model more smoothly. Typically, this nesting occurs through telescopic "multi" nesting approach tapering from an outer domain on the order of 1,500 km by 1,500 km to an inner domain that may be a 100 km by 100 km in extent. In this approach, the outer nest moves the lateral boundaries as far away from the desired model center of interest as possible so that tendencies from the external coarse-grain model pass gradually into the model domain. The

middle nest (or several) acts as an intermediate resolution nest(s) for ideal downscaling. Finally, the interior domain captures the domain of interest at the highest resolution desired. Such an approach is also called a "limited-area" mesoscale NWP configuration.

While the conservation-balance equations govern the transport of mass, energy, etc., throughout the model, other physics (often unresolved processes such as turbulence, whose effects must be estimated) are incorporated via "parameterizations" that capture effects of cloud cover, turbulence, solar radiation, etc. The collection of these conservation balance equations (partial differential equations that are linearized and solved numerically) along with the various parameterizations constitute the typical NWP model.

A single parameterization scheme will fall into one of 5 groups: 1) Microphysics, 2) Cumulus, 3) Radiation, 4) Planetary Boundary Layer, or 5) Surface effects that capture (respectively) physics such as moisture, clouds, solar radiation, turbulence, and land cover. The subgrid effects are a consequence of the interaction between representative schemes from each group during the simulation execution for the conservation principles to hold. Within each group there are a number of different approaches to handling that particular effect, each approach suitable for a range of conditions.

If one considers only the surface-layer and boundary-layer parameterizations, there are 15 combinations of these 2 physical schemes possible. If one takes all of the available combinations of physics options in WRF-ARW, there are more than 2 million possible combinations; thus, on a battlefield or in a remote deployment it is not possible to perform forecasts with all of these combinations to find the "single" set that best describes the local weather. Furthermore, each scheme or combination of schemes quite often only works best in certain environments. Given that we cannot know in advance where a brigade will be deployed, we seek a combination parameterization that works ideally well in all situations but practically well within a defined region.

If we consider our modeling system a "gray box", we can model the outputs (for example, forecast skill) as a statistical function of the initial conditions along with the internal physics that constitutes the "gray box". Consequently, our task becomes to appropriately sample that function in a manner that minimizes both the computational burden we face and allows us to maximize the amount of information we can extract from our set of model runs. Ideally, one should approach the selection of a candidate set of parameterizations via a method that leads to robust performance under a variety of weather conditions for a given domain—and one such approach is through the use of Design of Experiments techniques. Here

we intend a design point to be a single sample of the “gray-box simulation” whose value will be an output figure of merit such as forecast skill. Consequently, the design as a whole can be used to support identification of a suite of parameterization schemes that collectively produce a “skillful” forecast over a variety of conditions.

Appendix B. List of Confidence Index (CI) Predictors

This appendix appears in its original form, without editorial change.

CI v0.3b is thus based on 10 Rules which calculate a total of 18 predictor variables (features) which are used in the regression equation to calculate CI. Table 1 provides a descriptive list of the 18 predictor variables used in CI v0.3b.

Table List of predictors

	Predictor	Parent Component/System	Overview
1	GRAD_MIN	Predictor for the Gradient Rule (Rule 1)	Represents the lowest gradient value at 500 mb.
2	GRAD_MAX	Predictor for the Gradient Rule (Rule 1)	Represents the highest gradient value at 500 mb.
3	MAX_CLOSEDLO	Predictor for the Closed Low Rule (Rule 2)	Gives a numerical value for the evidence of existing closed low system in the AOI.
4	FIELD_DIFF_CON	Predictor for the Forecast Consistency Rule (Rule 3)	Numerical value to estimate the consistency between successive forecasts.
5	FIELD_DIFF_VER	Used for verification analysis (Rule 4)	Provides a numerical value as a means to validate the forecast against the analysis.
6	MAX_RV	Predictor for the New Closed Feature Analysis Rule (Rule 5)	Maximum relative vorticity at 500 mb
7	MAX_RV_SIZE	Predictor for the New Closed Feature Analysis Rule (Rule 5)	Measure of the size of largest region of significant relative vorticity
8	MAX_MAGGEO	Predictor for the Trough/Jets Analysis using Geostrophicnds at 500mb (Rule	Maximum magnitude of geostrophic wind at 500 mb
9	MAX_MAGGEO_SIZE	Predictor for the Trough/Jets Analysis using Geostrophic Winds at 500mb (Rule 6).	Measure of the size of largest region of significant geostrophic wind
10	SIG_MAGGEO_SIZE	Predictor for the Trough/Jets Analysis using Geostrophic Winds at 500mb (Rule 6).	Measure of the extent of all regions of significant geostrophic wind
11	MAX_MAGAGEO	Predictor for the Trough/Jets Analysis using Ageostrophic Winds at 500mb (Rule 7).	Maximum magnitude of *ageostrophic* wind at 500 mb
12	MAX_MAGAGEO_SIZE	Predictor for the Trough/Jets Analysis using Ageostrophic Winds at 500mb (Rule 7).	Measure of the size of largest region of significant *ageostrophic* wind
13	SIG_MAGAGEO_SIZE	Predictor for the Trough/Jets Analysis using Ageostrophic Winds at 500mb (Rule 7).	Measure of the extent of all regions of significant *ageostrophic* wind
14	CAPE_MAX	Predictor for Convection (Rule 8)	Maximum convective available potential energy
15	CIN_MAX	Predictor for Convection (Rule 8)	Maximum convective inhibition
16	DIV_MAX	Predictor for Divergence (Rule 9)	Maximum 500-mb wind divergence
17	DIV_MAX_SIZE	Predictor for Divergence (Rule 9)	Measure of the extent of all regions of significant 500-mb divergent wind
18	CONV_MAX	Predictor for Convergence (Rule 10)	Maximum 500-mb wind convergence
19	CONV_MAX_SIZE	Predictor for Convergence (Rule 10)	Measure of the extent of all regions of significant 500-mb convergence wind

List of Symbols, Abbreviations, and Acronyms

4D	4-dimensional
CI	confidence index
GFS	Global Forecast System
GIS	Geographic Information System
h	hour
hPa	hectopascal
MET	meteorological
METGM	Meteorological Gridded Message
NWP	Numerical Weather Prediction
PDF	probability distribution function
RMSE	root-mean-square error
WRF	Weather Research and Forecasting

1 (PDF)	DEFENSE TECHNICAL INFORMATION CTR DTIC OCA
5 (PDF)	DIRECTOR US ARMY RESEARCH LAB RDRL CIO LL IMAL HRA MAIL & RECORDS MGMT RDRL CIE JAMES L COGAN RDRL CIE M PATRICK A HAINES JEFFREY A SMITH
1 (PDF)	GOVT PRINTG OFC A MALHOTRA
3 (PDFs)	US ARMY RSRCH DEVEL AND ENGINEERING COMM ARMAMENT RSRCH DEVEL ENGINEERING CENTER WEAPONS AND SOFTWARE ENGINEERING CENTER FIRE CONTROL SYSTEMS AND TECH DIRECTORATE FIRING TABLES AND BALLIST DIV RDAR WSF T A J SOWA J A FONNER R M WELLS